

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/668,047A  
Source: 1FW16  
Date Processed by STIC: 2/24/05

***ENTERED***



IFW16

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005

TIME: 12:03:09

Input Set : A:\NEB-232.ST25.txt

Output Set: N:\CRF4\02242005\J668047A.raw

3 <110> APPLICANT: Lunnen, Keith  
 4       Davis, Theodore  
 5       Wilson, Geoffrey  
 7 <120> TITLE OF INVENTION: Method for Cloning and Expression of SbfI Restriction  
 8       Endonuclease and SbfI Methylase in E. coli  
 10 <130> FILE REFERENCE: NEB-232  
 12 <140> CURRENT APPLICATION NUMBER: 10/668,047A  
 13 <141> CURRENT FILING DATE: 2003-09-22  
 15 <150> PRIOR APPLICATION NUMBER: 10/668,047  
 16 <151> PRIOR FILING DATE: 2003-09-22  
 18 <160> NUMBER OF SEQ ID NOS: 57  
 20 <170> SOFTWARE: PatentIn version 3.2  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 1461  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: streptomyces species Bf-61  
 28 <220> FEATURE:  
 29 <221> NAME/KEY: CDS  
 30 <222> LOCATION: (1)..(1461)  
 32 <400> SEQUENCE: 1  
 33 gtc cat ccg atc gcc agc act gaa act cgc cgc caa gct gct ctc ggc       48  
 34 Met His Pro Ile Ala Ser Thr Glu Thr Arg Arg Gln Ala Ala Leu Gly  
 35 1               5               10               15  
 37 aaa ctg gac ccc act act caa gcg gtc cta ggg cag ttc ttc act ccc       96  
 38 Lys Leu Asp Pro Thr Thr Gln Ala Val Leu Gly Gln Phe Phe Thr Pro  
 39               20               25               30  
 41 atg aag gcc gcc acg ctg atg gct tca atg ctt cgg gtc gat gat ctc       144  
 42 Met Lys Ala Ala Thr Leu Met Ala Ser Met Leu Arg Val Asp Asp Leu  
 43               35               40               45  
 45 cgc gga acg gtg cgg gtc gac cca gga gct ggt gtc ggg tct ctg       192  
 46 Arg Gly Thr Val Arg Val Leu Asp Pro Gly Ala Gly Val Gly Ser Leu  
 47               50               55               60  
 49 acc gct gcc ctc gtc gat cgg ctg cat act gaa cgc ccc gac gtt gcg       240  
 50 Thr Ala Ala Leu Val Asp Arg Leu His Thr Glu Arg Pro Asp Val Ala  
 51 65               70               75               80  
 53 gtc cac gta gtt gcc gtg gaa acc gac ccc ttt gtc gtg cct tac ctg       288  
 54 Val His Val Val Ala Val Glu Thr Asp Pro Phe Val Val Pro Tyr Leu  
 55               85               90               95  
 57 cgc gcc acc ctg gag gaa tgt cgg aac gct tac ggc atc tcc tac gac       336  
 58 Arg Ala Thr Leu Glu Glu Cys Arg Asn Ala Tyr Gly Ile Ser Tyr Asp  
 59               100               105               110  
 61 cta gtc gag ggc gac tat ttg ctt aac caa ggg gcc aag ctg gat ggc       384  
 62 Leu Val Glu Gly Asp Tyr Leu Leu Asn Gln Gly Ala Lys Leu Asp Gly

P.6

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005  
TIME: 12:03:09

Input Set : A:\NEB-232.ST25.txt  
Output Set: N:\CRF4\02242005\J668047A.raw

63	115	120	125	
65	ccg ttc gat ctt gta att gct aat cct ccc tac gga aag ctt gct tca			432
66	Pro Phe Asp Leu Val Ile Ala Asn Pro Pro Tyr Gly Lys Leu Ala Ser			
67	130	135	140	
69	gat tcg ctg gcg cgg ctt gca acg aca gcg cgt gcc gtc gat gta ccg			480
70	Asp Ser Leu Ala Arg Leu Ala Thr Thr Ala Arg Ala Val Asp Val Pro			
71	145	150	155	160
73	aac gtt tac gtg gcc ttc tgg gtg cga gca gtc att tcg ctc aaa gag			528
74	Asn Val Tyr Val Ala Phe Trp Val Arg Ala Val Ile Ser Leu Lys Glu			
75	165	170	175	
77	cag ggg cgg ggg gtt ttc att gtt cct cga tct tgg gcg aac ggg cct			576
78	Gln Gly Arg Gly Val Phe Ile Val Pro Arg Ser Trp Ala Asn Gly Pro			
79	180	185	190	
81	tac tat cgt caa ttt cgc cat tgg ctg atg acc gcg gta agt ctc gat			624
82	Tyr Tyr Arg Gln Phe Arg His Trp Leu Met Thr Ala Val Ser Leu Asp			
83	195	200	205	
85	ata ctt cat gtg ttc gaa agt aga acc aaa gta ttt gcg gac acg aag			672
86	Ile Leu His Val Phe Glu Ser Arg Thr Lys Val Phe Ala Asp Thr Lys			
87	210	215	220	
89	gta aag caa gag aat gtc atc gtt gct ttc agt gtg agg ccg caa agc			720
90	Val Lys Gln Glu Asn Val Ile Val Ala Phe Ser Val Arg Pro Gln Ser			
91	225	230	235	240
93	tct agt gtg gtc ctt tct agg tcg gtc gca cat gga gaa gag tcg atc			768
94	Ser Ser Val Val Leu Ser Arg Ser Val Ala His Gly Glu Glu Ser Ile			
95	245	250	255	
97	gca agt tct gtg ccg ttt tct gcg ctt gtt cat gat gaa gac gat gac			816
98	Ala Ser Ser Val Pro Phe Ser Ala Leu Val His Asp Glu Asp Asp Asp			
99	260	265	270	
101	aaa atc gtg cac ttc gcg gaa agc gca tcg gtg ccc tcg gcg gcg agg			864
102	Lys Ile Val His Phe Ala Glu Ser Ala Ser Val Pro Ser Ala Ala Arg			
103	275	280	285	
105	ttt act ctc gct gat ctc ggc atc ggt gta agt acg gga aag gtt gtt			912
106	Phe Thr Leu Ala Asp Leu Gly Ile Gly Val Ser Thr Gly Lys Val Val			
107	290	295	300	
109	gat ttt cgc aat cgt cag tat ttg acc gat aac ctg gat gct tca ggc			960
110	Asp Phe Arg Asn Arg Gln Tyr Leu Thr Asp Asn Leu Asp Ala Ser Gly			
111	305	310	315	320
113	gtt gtg ccc atg gtt tat cag tca aac att cga tct ggt aaa att gat			1008
114	Val Val Pro Met Val Tyr Gln Ser Asn Ile Arg Ser Gly Lys Ile Asp			
115	325	330	335	
117	tgg cct cag gtg ggt gcg agg aag cct caa gga ttt gtt gcg gtc gaa			1056
118	Trp Pro Gln Val Gly Ala Arg Lys Pro Gln Gly Phe Val Ala Val Glu			
119	340	345	350	
121	gat gta gca cta cgt caa ctt ctc ccg caa ggg tcg tat gtt gtt gtg			1104
122	Asp Val Ala Leu Arg Gln Leu Leu Pro Gln Gly Ser Tyr Val Val Val			
123	355	360	365	
125	aaa cgg caa acg gcg aaa gag gac cgt cgt gtc atc gct gcg gtc			1152
126	Lys Arg Gln Thr Ala Lys Glu Asp Arg Arg Arg Val Ile Ala Ala Val			
127	370	375	380	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005  
TIME: 12:03:09

Input Set : A:\NEB-232.ST25.txt  
Output Set: N:\CRF4\02242005\J668047A.raw

129 tgg gac ggg gcc agc agg gtt gcg ctc gac aat aaa acg aac tat ttg	1200
130 Trp Asp Gly Ala Ser Arg Val Ala Leu Asp Asn Lys Thr Asn Tyr Leu	
131 385 390 395 400	
133 cat gaa tct caa cga ccg ctt gag aaa aat gtg gcc cgc ggc ctc atg	1248
134 His Glu Ser Gln Arg Pro Leu Glu Lys Asn Val Ala Arg Gly Leu Met	
135 405 410 415	
137 ctt tgg ttg aac tcg act gtg ttg gat cag tat ttc cga gcc ttt tcc	1296
138 Leu Trp Leu Asn Ser Thr Val Leu Asp Gln Tyr Phe Arg Ala Phe Ser	
139 420 425 430	
141 ggg cat acc cag gtg aac gct ggc gat cta cgc cgg ctt ccg ttc ctc	1344
142 Gly His Thr Gln Val Asn Ala Gly Asp Leu Arg Arg Leu Pro Phe Leu	
143 435 440 445	
145 tgt cgc gag gac cta att ctt ctc gct aag gtc gtt ccc gat ggc ctg	1392
146 Cys Arg Glu Asp Leu Ile Leu Ala Lys Val Val Pro Asp Gly Leu	
147 450 455 460	
149 cct gat cag gag acg ttg gat gcc gtg gtg gcc aga ctc ttc tgt gag	1440
150 Pro Asp Gln Glu Thr Leu Asp Ala Val Val Ala Arg Leu Phe Cys Glu	
151 465 470 475 480	
153 att ccg gaa tct gcc tcg tga	1461
154 Ile Pro Glu Ser Ala Ser	
155 485	
158 <210> SEQ ID NO: 2	
159 <211> LENGTH: 486	
160 <212> TYPE: PRT	
161 <213> ORGANISM: streptomyces species Bf-61	
163 <400> SEQUENCE: 2	
165 Met His Pro Ile Ala Ser Thr Glu Thr Arg Arg Gln Ala Ala Leu Gly	
166 1 5 10 15	
169 Lys Leu Asp Pro Thr Thr Gln Ala Val Leu Gly Gln Phe Phe Thr Pro	
170 20 25 30	
173 Met Lys Ala Ala Thr Leu Met Ala Ser Met Leu Arg Val Asp Asp Leu	
174 35 40 45	
177 Arg Gly Thr Val Arg Val Leu Asp Pro Gly Ala Gly Val Gly Ser Leu	
178 50 55 60	
181 Thr Ala Ala Leu Val Asp Arg Leu His Thr Glu Arg Pro Asp Val Ala	
182 65 70 75 80	
185 Val His Val Val Ala Val Glu Thr Asp Pro Phe Val Val Pro Tyr Leu	
186 85 90 95	
189 Arg Ala Thr Leu Glu Glu Cys Arg Asn Ala Tyr Gly Ile Ser Tyr Asp	
190 100 105 110	
193 Leu Val Glu Gly Asp Tyr Leu Leu Asn Gln Gly Ala Lys Leu Asp Gly	
194 115 120 125	
197 Pro Phe Asp Leu Val Ile Ala Asn Pro Pro Tyr Gly Lys Leu Ala Ser	
198 130 135 140	
201 Asp Ser Leu Ala Arg Leu Ala Thr Thr Ala Arg Ala Val Asp Val Pro	
202 145 150 155 160	
205 Asn Val Tyr Val Ala Phe Trp Val Arg Ala Val Ile Ser Leu Lys Glu	
206 165 170 175	
209 Gln Gly Arg Gly Val Phe Ile Val Pro Arg Ser Trp Ala Asn Gly Pro	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005  
TIME: 12:03:09

Input Set : A:\NEB-232.ST25.txt  
Output Set: N:\CRF4\02242005\J668047A.raw

210	180	185	190
213	Tyr Tyr Arg Gln Phe Arg His Trp Leu Met Thr Ala Val Ser Leu Asp		
214	195	200	205
217	Ile Leu His Val Phe Glu Ser Arg Thr Lys Val Phe Ala Asp Thr Lys		
218	210	215	220
221	Val Lys Gln Glu Asn Val Ile Val Ala Phe Ser Val Arg Pro Gln Ser		
222	225	230	235
225	Ser Ser Val Val Leu Ser Arg Ser Val Ala His Gly Glu Glu Ser Ile		
226	245	250	255
229	Ala Ser Ser Val Pro Phe Ser Ala Leu Val His Asp Glu Asp Asp Asp		
230	260	265	270
233	Lys Ile Val His Phe Ala Glu Ser Ala Ser Val Pro Ser Ala Ala Arg		
234	275	280	285
237	Phe Thr Leu Ala Asp Leu Gly Ile Gly Val Ser Thr Gly Lys Val Val		
238	290	295	300
241	Asp Phe Arg Asn Arg Gln Tyr Leu Thr Asp Asn Leu Asp Ala Ser Gly		
242	305	310	315
245	320		
246	Val Val Pro Met Val Tyr Gln Ser Asn Ile Arg Ser Gly Lys Ile Asp		
249	325	330	335
250	Trp Pro Gln Val Gly Ala Arg Lys Pro Gln Gly Phe Val Ala Val Glu		
253	340	345	350
254	Asp Val Ala Leu Arg Gln Leu Leu Pro Gln Gly Ser Tyr Val Val Val		
257	355	360	365
258	370	375	380
261	Lys Arg Gln Thr Ala Lys Glu Asp Arg Arg Arg Val Ile Ala Ala Val		
262	385	390	395
265	400		
266	Trp Asp Gly Ala Ser Arg Val Ala Leu Asp Asn Lys Thr Asn Tyr Leu		
269	405	410	415
270	Leu Trp Leu Asn Ser Thr Val Leu Asp Gln Tyr Phe Arg Ala Phe Ser		
273	420	425	430
274	Gly His Thr Gln Val Asn Ala Gly Asp Leu Arg Arg Leu Pro Phe Leu		
277	435	440	445
278	Cys Arg Glu Asp Leu Ile Leu Ala Lys Val Val Pro Asp Gly Leu		
281	450	455	460
282	Pro Asp Gln Glu Thr Leu Asp Ala Val Val Ala Arg Leu Phe Cys Glu		
285	465	470	475
286	Ile Pro Glu Ser Ala Ser		
289	485		
290	<210> SEQ ID NO: 3		
291	<211> LENGTH: 972		
292	<212> TYPE: DNA		
295	<213> ORGANISM: streptomyces species Bf-61		
296	<220> FEATURE:		
297	<221> NAME/KEY: CDS		
299	<222> LOCATION: (1)..(972)		
300	<400> SEQUENCE: 3		
301	gtg aac agc agt gac ggc atc gac gga acg gta gcg agc atc gat act		48
	Met Asn Ser Ser Asp Gly Ile Asp Gly Thr Val Ala Ser Ile Asp Thr		

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005  
TIME: 12:03:09

Input Set : A:\NEB-232.ST25.txt  
Output Set: N:\CRF4\02242005\J668047A.raw

302 1	5	10	15															
304	gcg	cg	ctg	ctg cta aag cgt ttt ggg ttt gac gcg caa cga tat aac	96													
305	Ala	Arg	Ala	Leu	Leu	Lys	Arg	Phe	Gly	Phe	Asp	Ala	Gln	Arg	Tyr	Asn		
306																	30	
308	gtc	cgt	agc	gct	gtg	aca	ttg	ctc	g	ctt	gcc	ggt	ttg	aag	ccg	gga	144	
309	Val	Arg	Ser	Ala	Val	Thr	Leu	Leu	Ala	Leu	Ala	Gly	Leu	Lys	Pro	Gly		
310																	45	
312	gat	cgc	tgg	gtt	gac	tcg	acc	act	cca	cgc	ctt	ggc	gtt	cag	aag	atc	192	
313	Asp	Arg	Trp	Val	Asp	Ser	Thr	Thr	Pro	Arg	Leu	Gly	Val	Gln	Lys	Ile		
314																	60	
316	atg	gac	tgg	tcc	ggc	gag	cat	tgg	gcc	aag	ccg	tac	gcc	acc	gga	agt	240	
317	Met	Asp	Trp	Ser	Gly	Glu	His	Trp	Ala	Lys	Pro	Tyr	Ala	Thr	Gly	Ser		
318	65																80	
320	cga	gaa	gat	ttc	cgt	aag	aag	acg	ctt	cg	cag	tgg	gtt	gat	aat	ggc	288	
321	Arg	Glu	Asp	Phe	Arg	Lys	Lys	Thr	Leu	Arg	Gln	Trp	Val	Asp	Asn	Gly		
322																	95	
324	ttc	gcc	gta	ctt	aat	g	cg	gac	aat	tta	aa	atc	gcc	acg	aa	tcg	336	
325	Phe	Ala	Val	Leu	Asn	Ala	Asp	Asn	Leu	Asn	Ile	Ala	Thr	Asn	Ser	Gln		
326																	110	
328	ctc	aac	gag	tac	tgc	ttg	tct	gac	gaa	gca	tta	cag	g	ct	agg	gca	384	
329	Leu	Asn	Glu	Tyr	Cys	Leu	Ser	Asp	Glu	Ala	Leu	Gln	Ala	Leu	Arg	Ala		
330																	125	
332	tat	gga	acg	gaa	ggc	ttc	gag	gaa	tct	ctt	gta	gtc	ttt	ctt	gat	aat	ggc	432
333	Tyr	Gly	Thr	Glu	Gly	Phe	Glu	Glu	Ser	Leu	Val	Val	Phe	Leu	Asp	Glu		
334																	140	
336	gca	tcg	aag	g	cg	gtt	aaa	g	cg	g	gaa	g	ctc	cag	g	ct	480	
337	Ala	Ser	Lys	Ala	Val	Lys	Ala	Arg	Ala	Glu	Ala	Leu	Gln	Ala	Ala	Met		
338	145																160	
340	atc	tct	gtc	gat	ctc	cct	gg	ggc	gag	gaa	ttt	ctg	ctc	tcg	cct	gcc	528	
341	Ile	Ser	Val	Asp	Leu	Pro	Gly	Gly	Glu	Glu	Phe	Leu	Leu	Ser	Pro	Ala		
342																	175	
344	ggg	cag	aat	cca	ttg	ctg	aag	aag	atg	gtc	gaa	gag	ttt	gtg	ccg	cg	576	
345	Gly	Gln	Asn	Pro	Leu	Leu	Lys	Lys	Met	Val	Glu	Glu	Phe	Val	Pro	Arg		
346																	190	
348	ttt	gca	cct	cg	tcg	ac	gt	ct	tc	ct	gg	gat	act	cgt	gg	aag	624	
349	Phe	Ala	Pro	Arg	Ser	Thr	Val	Leu	Tyr	Leu	Gly	Asp	Thr	Arg	Gly	Lys		
350																	205	
352	cat	tcc	cta	ttc	gaa	cg	gag	atc	ttt	gaa	gag	gt	ct	gg	ct	act	672	
353	His	Ser	Leu	Phe	Glu	Arg	Glu	Ile	Phe	Glu	Glu	Val	Leu	Gly	Leu	Thr		
354																	220	
356	ttc	gac	ccc	cat	gg	cg	atg	ccg	gac	ctt	att	ctc	cat	gac	gaa	gtt	720	
357	Phe	Asp	Pro	His	Gly	Arg	Met	Pro	Asp	Leu	Ile	Leu	His	Asp	Glu	Val		
358	225																240	
360	cgt	ggg	tgg	ttt	tc	tt	atg	gag	ggc	gt	aaa	agt	aaa	ggt	ccg	ttt	768	
361	Arg	Gly	Trp	Leu	Phe	Leu	Met	Glu	Ala	Val	Lys	Ser	Lys	Gly	Pro	Phe		
362																	255	
364	gat	gag	gag	cg	cat	cg	ag	cg	caa	gag	ct	ttc	gtt	aca	cct	tca	816	
365	Asp	Glu	Glu	Arg	His	Arg	Ser	Leu	Gln	Glu	Leu	Phe	Val	Thr	Pro	Ser		
366																	270	

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005  
TIME: 12:03:10

Input Set : A:\NEB-232.ST25.txt  
Output Set: N:\CRF4\02242005\J668047A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; N Pos. 3,12

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/668,047A

DATE: 02/24/2005

TIME: 12:03:10

Input Set : A:\NEB-232.ST25.txt

Output Set: N:\CRF4\02242005\J668047A.raw

L:554 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0